

Analisis Regresi Multivariat Pada Nilai Derajat Kesehatan Penduduk

Adela Ari Artha

Jurusan Matematika, FMIPA, Universitas Negeri Surabaya

e-mail : adelaartha@mhs.unesa.ac.id

Abstrak

Seiring dengan perkembangan jaman yang sangat cepat, sebuah penelitian menjadi aspek penting dalam bidang keilmuan. Variable-variable yang digunakan dalam penelitian pun menjadi tidak terbatas. Seperti contoh dalam penelitian ini akan dilakukan prediksi derajat kesehatan penduduk. Dalam melakukan prediksi derajat kesehatan penduduk tidak hanya melibatkan satu input dan satu output melainkan menggunakan empat input dan dua output. Variable input yang digunakan yaitu Presentase Penduduk dengan akses sanitasi layak, Presentase rumah tangga yang sehat, Presentase peserta KB Baru, dan Presentase bayi yang mendapat ASI eksklusif. Sedangkan variable output yang digunakan untuk menentukan derajat kesehatan penduduk yaitu angka kematian bayi (AKB) dan angka gizi buruk. Metode yang digunakan untuk memprediksi hubungan antara variable input dan output adalah metode regresi multivariat secara manual. Sehingga diperoleh kesimpulan bahwa variabel independen / variable input (x 's) mempengaruhi variabel dependen / variable output (y 's) berdasarkan hasil dari nilai Wilks Lambda yang diperoleh.

Kata kunci: Regresi multivariat, derajat kesehatan penduduk, Wilks Lambda

Abstract

Along with the rapid development, a study has become an important aspect of the scientific field. The variables used in the study became unlimited. As an example in this study will predict the health status of the population. In predicting the health status of the population it does not only involve one input and one output but also uses four inputs and two outputs. Input variables used are Percentage of population with access to proper sanitation, Percentage of healthy households, Percentage of new family planning participants, and Percentage of infants who were exclusively breastfed. While the output variables that are used to determine the health status of the population, namely the infant mortality rate (IMR) and malnutrition. The method used to predict the correlation between input and output variable is manually multivariate regression method. So we concluded that the independent variable/input variables (x 's) affect the dependent variable/output variable (y 's) based on the results of Wilks Lambda value obtained

Keywords: Multivariate regression, population health status, Wilks Lambda

1. PENDAHULUAN

Derajat kesehatan penduduk merupakan aspek penting untuk mengetahui kesejahteraan penduduk dilihat dari bagaimana kualitas fisik kesehatan penduduk. Untuk memprediksi derajat kesehatan penduduk maka digunakan beberapa variable yang dibutuhkan untuk mengolah data. Derajat kesehatan penduduk dapat dilihat melalui dua factor yaitu angka kematian bayi dan angka gizi buruk yang diperoleh. Dua factor ini yang nantinya dianggap sebagai variable output / variable dependen (y). Sedangkan angka kematian bayi dan angka gizi buruk dapat dipengaruhi oleh beberapa factor berikut yaitu Presentase Penduduk dengan akses sanitasi layak, Presentase rumah tangga yang sehat, Presentase peserta KB Baru, dan Presentase bayi yang mendapat ASI eksklusif. Variable ini yang nantinya dianggap sebagai variable input / variable independen (x). Sehingga variable output yang digunakan yaitu angka kematian bayi (y_1) dan angka gizi buruk (y_2)

dan variable input yang digunakan yaitu presentase penduduk dengan akses sanitasi layak (x_1), presentase rumah tangga yang sehat (x_2), presentase peserta KB baru (x_3), dan presentase bayi yang mendapat ASI eksklusif (x_4). Metode yang digunakan dalam penelitian ini yaitu menggunakan metode regresi multivariate. Prediksi regresi multivariate adalah prediksi statistika yang dikenakan pada data yang terdiri dari banyak variabel dan antar variabel saling berkorelasi. Data Multivariat tidak hanya terdiri atas satu variabel saja melainkan dapat terdiri atas lebih dari satu variabel. Misal data dari n pengamatan pada p variable. Sehingga dapat disusun matriks dengan n baris dan p kolom, yang dinotasikan X . Tujuan dari penelitian ini yaitu untuk mengetahui apakah ada perbedaan signifikan yang mempengaruhi penelitian dalam memprediksi hubungan antara variable input (x) dan variable output (y).

2. KAJIAN TEORI

Model Regresi Multivariat

Pada regresi multivariat ini melibatkan beberapa y 's yang dipertimbangkan dengan baik kesesuaiannya pada setiap himpunan pada x 's. Setiap y_1, y_2, \dots, y_p diprediksi dengan semua x_1, x_2, \dots, x_q .

Misal terdapat n observasi, \mathbf{Y} merupakan variabel dependent sebanyak p dan \mathbf{X} merupakan variabel independen sebanyak q . Maka matriks \mathbf{Y} dan \mathbf{X} dapat ditulis sebagai berikut (Rencher, 2001):

$$\mathbf{Y} = \begin{bmatrix} y_{11} & y_{12} & \dots & y_{1p} \\ y_{21} & y_{22} & \dots & y_{2p} \\ \vdots & \vdots & \dots & \vdots \\ y_{n1} & y_{n2} & \dots & y_{np} \end{bmatrix} = \begin{pmatrix} y'_1 \\ y'_2 \\ \vdots \\ y'_n \end{pmatrix}$$

$$\mathbf{X} = \begin{bmatrix} 1 & x_{11} & \dots & x_{1q} \\ 1 & x_{21} & \dots & x_{2q} \\ \vdots & \vdots & \dots & \vdots \\ 1 & x_{n1} & \dots & x_{nq} \end{bmatrix}$$

Karena setiap p y 's bergantung pada x 's di setiap cara, setiap kolom pada \mathbf{Y} akan berbeda β 's. Maka kolom pada β 's untuk setiap matrik \mathbf{Y} dan bentuk kolom matrik $\mathbf{B} = (\beta_1, \beta_2, \dots, \beta_p)$. Sehingga model regresi multivariat adalah sebagai berikut (Edi, 2014):

$$\mathbf{Y} = \mathbf{XB} + \mathbf{Z}$$

Maka persamaan regresi multivariat dapat ditulis sebagai berikut:

$$\begin{bmatrix} y_{11} & y_{12} & \dots & y_{1p} \\ y_{21} & y_{22} & \dots & y_{2p} \\ \vdots & \vdots & \dots & \vdots \\ y_{n1} & y_{n2} & \dots & y_{np} \end{bmatrix} = \begin{bmatrix} 1 & x_{11} & \dots & x_{1q} \\ 1 & x_{21} & \dots & x_{2q} \\ \vdots & \vdots & \dots & \vdots \\ 1 & x_{n1} & \dots & x_{nq} \end{bmatrix} \begin{bmatrix} \beta_{01} & \beta_{02} & \dots & \beta_{0p} \\ \beta_{11} & \beta_{12} & \dots & \beta_{1p} \\ \vdots & \vdots & \dots & \vdots \\ \beta_{p1} & \beta_{p2} & \dots & \beta_{pq} \end{bmatrix} + \begin{bmatrix} \varepsilon_{11} & \varepsilon_{12} & \dots & \varepsilon_{1q} \\ \varepsilon_{21} & \varepsilon_{22} & \dots & \varepsilon_{2q} \\ \vdots & \vdots & \dots & \vdots \\ \varepsilon_{n1} & \varepsilon_{n2} & \dots & \varepsilon_{nq} \end{bmatrix}$$

Keterangan:

\mathbf{Y} = Matriks peubah dependen berukuran $n \times q$

\mathbf{X} = Matriks peubah independen berukuran $n \times (p + 1)$

β = Matriks parameter regresi dengan ukuran $(p + 1) \times q$

ε = Matriks residual berukuran $n \times q$

Terbentuknya matriks β dapat dicari dengan tahapan seperti dibawah ini:

Diilustrasikan model multivariat dengan $p = 2, q = 3$ dan $n = 5$, dengan p adalah variabel dependen dan q adalah variabel independen, sehingga:

$$\begin{bmatrix} y_{11} & y_{12} \\ y_{21} & y_{22} \\ y_{31} & y_{32} \\ y_{41} & y_{42} \\ y_{51} & y_{52} \end{bmatrix} = \begin{bmatrix} 1 & x_{11} & x_{12} & x_{13} \\ 1 & x_{21} & x_{22} & x_{23} \\ 1 & x_{31} & x_{32} & x_{33} \\ 1 & x_{41} & x_{42} & x_{43} \\ 1 & x_{51} & x_{52} & x_{53} \end{bmatrix} \begin{bmatrix} \beta_{11} & \beta_{12} \\ \beta_{21} & \beta_{22} \\ \beta_{31} & \beta_{32} \\ \beta_{41} & \beta_{42} \\ \beta_{51} & \beta_{52} \end{bmatrix} + \begin{bmatrix} \varepsilon_{11} & \varepsilon_{12} \\ \varepsilon_{21} & \varepsilon_{22} \\ \varepsilon_{31} & \varepsilon_{32} \\ \varepsilon_{41} & \varepsilon_{42} \\ \varepsilon_{51} & \varepsilon_{52} \end{bmatrix}$$

Pada matriks \mathbf{X} diberi nilai dummy yaitu 1, sebagai konstanta pada variabel independen.

Estimasi Least Square

Untuk estimasi parameter dengan menggunakan Least square dapat diselesaikan sebagai berikut (Johnson, 2007):

$$\hat{\mathbf{B}} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Y}$$

$\hat{\mathbf{B}}$ merupakan estimasi least square.

Matriks parameter regresi $\hat{\mathbf{B}}$ didapatkan dari hasil perkalian matriks \mathbf{Y} dan matriks \mathbf{X} dengan masing-masing nilai $\mathbf{Y} = 2$, $\mathbf{X} = 3$, dan $n = 5$ sebagai berikut:

$$\begin{aligned} \hat{\mathbf{B}} &= ((4x5)(5x4))^{-1}((4x5)(5x2)) \\ &= (4x4)^{-1}(4x2) \\ &= (4x2) \end{aligned}$$

sehingga didapatkan nilai matriks $\hat{\mathbf{B}}$ adalah sebesar $(4x2)$, dengan bentuk matriks pada $\hat{\mathbf{B}}$ sebagai berikut:

$$\begin{pmatrix} \beta_{11} & \beta_{12} \\ \beta_{21} & \beta_{22} \\ \beta_{31} & \beta_{32} \\ \beta_{41} & \beta_{42} \end{pmatrix}$$

Uji Hipotesis

Uji hipotesis adalah uji yang dilakukan untuk semua koefisien regresi β dan \mathbf{B}_1 terhadap \mathbf{Y} . Hipotesis yang digunakan yaitu sebagai berikut:

$$H_0: \mathbf{B}_1 = \mathbf{0}$$

dimana \mathbf{B}_1 masuk kedalam semua baris \mathbf{B} yang pertama:

$$\mathbf{B} = \begin{bmatrix} \beta'_0 \\ \mathbf{B}_1 \end{bmatrix} = \begin{bmatrix} \beta_{01} & \beta_{02} & \dots & \beta_{0p} \\ \beta_{11} & \beta_{12} & \dots & \beta_{1p} \\ \vdots & \vdots & \dots & \vdots \\ \beta_{q1} & \beta_{q2} & \dots & \beta_{qp} \end{bmatrix}$$

β'_0 tidak dimasukkan kedalam hipotesis karena akan mengakibatkan semua nilai pada y 's memiliki nilai nol (Morrison, 2005).

$$H_1: \mathbf{B}_1 \neq \mathbf{0}$$

$H_1: \mathbf{B}_1 \neq \mathbf{0}$ untuk mengetahui bahwa $\beta_{jk} \neq 0$, dengan $j = 1, 2, \dots, q; k = 1, 2, \dots, p$.

Sedangkan statistik uji yang digunakan adalah Wilk's lamda:

$$\Lambda = \frac{|\mathbf{E}|}{|\mathbf{E} + \mathbf{H}|} = \frac{|\mathbf{Y}'\mathbf{Y} - \hat{\mathbf{B}}'\mathbf{X}'\mathbf{Y}|}{|\mathbf{Y}'\mathbf{Y} - n\bar{y}\bar{y}'|}$$

Tolak H_0 jika $\Lambda_{hitung} \leq \Lambda_{\alpha, p, q, (n-q-1)}$. Nilai $\Lambda_{\alpha, p, q, (n-q-1)}$ merupakan nilai kritis dari tabel Wilk's Lamda.

3. METODE

Data yang digunakan dalam penelitian ini adalah data sekunder yang diperoleh dari data dan informasi profil kesehatan Indonesia 2015 yang diperoleh dari www.kemkes.go.id. Data terdiri dari 35 observasi yang merupakan 35 kabupaten/kota yang berada di Jawa Tengah. Data akan dianalisis menggunakan regresi multivariate

4. PEMBAHASAN

Untuk mengetahui derajat kesehatan penduduk, digunakan data sebagai berikut :

Tabel data informasi profil kesehatan

Obs	Y1	Y2	X1	X2	X3	X4
1	207	76	81,6	72,46	16,1	86,3
2	243	57	58,8	79,73	12,9	50,4
3	149	8	61,1	75,73	11,7	67,6
4	209	23	64,8	67,35	11,6	66,9
5	201	13	82,2	78,32	11	81,4
6	105	35	72,6	76,86	9,7	85
7	126	20	62,7	76,49	14,1	76,2
8	131	32	75,4	66,43	9,4	76
9	127	6	90	81,47	10,6	52,6
10	220	13	76,5	96,43	11,6	75,4
11	125	12	93,8	73,93	12,5	63,4
12	104	48	88	89,31	12,2	58,4
13	158	9	88,3	83,95	13,5	60,3
14	129	6	81,7	83,31	14,5	66,9
15	384	26	94,9	88,03	14	67,7
16	169	55	75,1	78,3	12	73,8
17	134	31	63	76,95	12,3	63,9
18	167	33	99,8	72,74	11,1	67
19	152	16	77	62,28	11,8	13,1
20	134	45	69,4	75,84	13,7	73,3
21	149	13	78,2	81,41	14,2	47,5
22	158	26	85,4	64,42	12,2	6,7
23	184	27	85,2	79,23	10,8	83,7
24	160	28	90	59,87	11,3	71,5
25	169	27	68,5	80,08	14,1	41,8
26	126	46	85,6	76,21	10	42,9
27	201	5	72,4	74,18	12,5	57,9
28	263	57	56,5	77,73	11,5	33,4
29	321	82	69,1	55,89	15,1	67,5
30	25	9	79,1	95,93	16,9	48,3
31	80	0	67,5	94,71	9,1	52,4
32	35	3	85,7	82,85	10,2	44,5
33	229	13	90	90,94	9,8	57,3
34	58	7	74	93,37	15	59,5
35	39	15	62	94,26	13,8	68

Dimana masing-masing variable di definisikan sebagai berikut :

Variabel	Deskripsi
Y ₁	Angka kematian Bayi (AKB)
Y ₂	Gizi Buruk
X ₁	Presentase Penduduk dengan akses sanitasi layak
X ₂	Presentase rumah tangga yang sehat
X ₃	Presentase peserta KB Baru
X ₄	Presentase bayi yang mendapat ASI eksklusif

Sehingga dapat diketahui :

$$n = 35 \quad p=2 \quad q=4$$

Dalam langkah pertama akan ditentukan model persamaan regresi :

$$Y = XB + Z$$

Untuk membentuk model dari persamaan regresi tersebut maka dibutuhkan parameter β yang diestimasi menggunakan metode estimasi parameter *least square* sebagai berikut :

$$\hat{B} = (X'X)^{-1}X'Y$$

Sehingga diperoleh estimasi parameter berikut :

$$\hat{B} = \begin{bmatrix} 252,8813127 & 82,60137 \\ 0,28601224 & -0,26797 \\ -2,421452184 & -0,95565 \\ 3,548025167 & 2,177753 \\ 0,515327341 & 0,212587 \end{bmatrix}$$

Setelah didapatkan matriks parameter \hat{B} kemudian dilakukan penentuan hipotesis untuk mengambil keputusan sebagai berikut

$$H_0: B_1 = 0$$

$$H_1: \text{paling sedikit ada satu } B_1 \neq 0$$

Kemudian dicari nilai Wilks Lambda untuk mengambil keputusan dengan rumus :

$$\Lambda = \frac{|E|}{|E + H|} = \frac{|Y'Y - \hat{B}'X'Y|}{|Y'Y - n\bar{y}\bar{y}'|}$$

Dan diperoleh hasil nilai Wilks Lambda yaitu ::

$$\Lambda = \frac{|E|}{|E + H|} = \frac{1508917269}{2216999327} = 0,6806124$$

sedangkan nilai Wilks Lambda berdasarkan table diperoleh sebesar 0,736. Dari perolehan nilai Wilks Lambda tersebut maka dapat disimpulkan jika

$$\Lambda_{hitung} < \Lambda_{\alpha;p;q;n-q-1}$$

maka tolak H_0 (Rita, 2015). Λ_{hitung} yang diperoleh yaitu sebesar 0,6806124 dan $\Lambda_{0,05;2;4;30}$ yaitu sebesar 0,736 sehingga didapat kesimpulan bahwa tolak H_0 dan dapat ditarik kesimpulan bahwa variabel independen (x 's) mempengaruhi variabel dependen (y 's).

5. PENUTUP

Simpulan

Dari penelitian ini dapat disimpulkan bahwa variable independen (x 's) yang terdiri dari presentase penduduk dengan akses sanitasi layak, presentase rumah tangga yang sehat, presentase peserta KB baru, dan presentasi bayi yang mendapat ASI eksklusif mempengaruhi besarnya derajat kesehatan penduduk berdasarkan angka kematian bayi dan angka gizi buruk.

Saran

Untuk penelitian selanjutnya dapat digunakan metode lain yang dapat menjelaskan hubungan antar beberapa variabel lebih baik.

DAFTAR PUSTAKA

- Edi, S. d. (2014). *Analisis Regresi Multivariat Pada Intention To Buy Berdasarkan Efektivitas Pemasaran*. IENACO: Jurnal Seminar Nasional.
- Johnson, R. &. (2007). *Applied Multivariat Statistical Analysis*. New Jersey: Prentice Hall.
- Morrison, D. (2005). *Multivariate Statistical Methods, Fourth Edition*. Pennsylvania: The Wharton School University of Pennsylvania.
- Rencher, A. (2001). *Methods of Multivariate Analysis*. USA: Wiley Interscience.
- Rita, R. V. (2015). Penerapan Regresi Linier Multivariat pada Distribusi Ujian Nasional tahun 2014. *Jurnal Gaussian Vol 4, No 3*, 2339-2541.

